

# Talen en Automaten

Test 2, Thu 21<sup>st</sup> Jan, 2016

This test consists of **four** exercises over **6 pages**. Explain your approach. You can score a maximum of 100 points, and each question indicates how many points it is worth. The test is closed book. You are NOT allowed to use a calculator, a computer or a mobile phone. You may answer in Dutch or in English. Please write clearly, and do not forget to put on each page: your name and your student number.

**Notation** Throughout the test, we denote for any alphabet  $A$  and  $a \in A$  by  $|w|_a$  the number of  $a$ 's in the word  $w \in A^*$ , as it was introduced in the lecture.

**Write the answer to each exercise on a separate sheet!**

## 1 Non-Regular Languages

**Write your answers on a separate sheet**

Let  $A = \{a, b\}$ .

- a) We define the language  $L$  to be

$$L = \{wb^n \mid w \in A^*, |w| = n\}.$$

Show that  $L$  is not regular.

**(5pt)**

**Solution:** .....

Assume that  $L$  is regular and let  $p > 0$  be the pumping length which we get from the pumping lemma (PL). Take  $w = a^p b^p \in L$ , which can be divided by the PL into  $w = xyz$  with  $|y| \geq 1$ ,  $|xy| \leq p$  such that  $xy^i z \in L$  for all  $i \in \mathbb{N}$ . By the above constraints we immediately get that  $y = a^k$  with  $k > 0$ . Take  $i = 0$ . So  $xy^0 z = a^{p-k} b^p \in L$ . But  $p - k \neq p$ , so  $xy^0 z \notin L$ . Contradiction, hence  $L$  is not regular.

Students may use the following (incorrect) argument “ $L \cap \mathcal{L}(a^* b^*) = \{a^n b^n \mid n \in \mathbb{N}\}$ , which is not regular. Thus  $L$  cannot be regular.” However:  $L \cap \mathcal{L}(a^* b^*) = \{a^m b^n \mid n \geq m \text{ and } n + m \text{ is even}\}$ , so this argument is incorrect.  $\square$

- b) Show that the language  $L = \{w \in A^* \mid |w|_a = |w|_b\}$  is not regular, **(10pt)** using the Pumping Lemma.

**Solution:** .....

Assume that  $L$  is regular and let  $p > 0$  be the pumping length which we get from the pumping lemma (PL). Take  $w = a^p b^p \in L$ , which can be divided by the PL into  $w = xyz$  with  $|y| \geq 1$ ,  $|xy| \leq p$  such that  $xy^i z \in L$  for all  $i \in \mathbb{N}$ . By the above constraints we immediately get that  $y = a^k$  with  $k > 0$ . Take  $i = 0$  (or  $i = 2$  also works). So  $xy^0 z = a^{p-k} b^p \in L$ . But  $p - k \neq p$ , so  $xy^0 z \notin L$ . Contradiction, hence  $L$  is not regular.  $\square$

## 2 Context Free Grammars

### Write your answers on a separate sheet

Fix  $A = \{a, b\}$  for this exercise.

a) Let  $L$  be the language over  $A$  given by  $L = \{a^n b^k a^m \mid k = n + m\}$ .

i) Construct a CFG  $G$  such that  $\mathcal{L}(G) = L$ . (10pt)

**Solution:** .....

$G$  is given by the productions

$$\begin{aligned} S &\longrightarrow LR \\ L &\longrightarrow aLb \mid \lambda \\ R &\longrightarrow bRa \mid \lambda \end{aligned}$$

having non-terminals  $\{S, L, R\}$  and start symbol  $S$ . □

ii) Give a derivation for the word  $aabbba \in L$ . (5pt)

**Solution:** .....

$$\begin{aligned} S &\Rightarrow LR \Rightarrow aLbR \Rightarrow aaLbbR \Rightarrow aabbR \\ &\Rightarrow aabbbRa \Rightarrow aabbba \end{aligned}$$

□

iii) Show that the word  $aba$  is not generated. (5pt)

**Solution:** .....

The only possible way to get the first  $ab$  is by

$$S \Rightarrow LR \Rightarrow aLbR \Rightarrow abR$$

but then from  $R$  we can only get to  $\lambda$  or to  $b\dots$ , so we cannot generate  $aba$ .

**Alternative: Full Case Exploration** Starting from  $S$ , we have the following possible derivations.

$$S \Rightarrow LR \Rightarrow \{aLbR, LbRa, R, L\}$$

For these, in turn, we have

•

$$aLbR \Rightarrow \{abR, aaLbbR, aLb, aLbbRa\}$$

Thus in each case either too little a's or too many b's are produced.

•

$$LbRa \Rightarrow \{bRa, aLbbRa, Lba, LbbRaa\}$$

Similar

- $R$  cannot produce a in front of b
- $L$  cannot produce b in front of a

Thus in neither of these cases we can derive  $aba$ , so it is not in the language generated by the grammar.  $\square$

b) Let  $G$  be the following CFG over  $A$ .

$$\begin{aligned} S &\longrightarrow US \mid \lambda \\ U &\longrightarrow aa \mid ab \mid bb \mid ba \end{aligned}$$

i) Give a precise description of  $\mathcal{L}(G)$  using set notation. (5pt)

**Solution:** .....

$$\begin{aligned} \mathcal{L}(G) &= \{w \in A^* \mid |w| \text{ even}\} \\ &= \{w_1 \cdots w_n \mid n \in \mathbb{N}, w_i \in \{aa, ab, bb, ba\}\} \\ &= \mathcal{L}((aa + ab + bb + ba)^*) \end{aligned}$$

$\square$

ii) Is  $\mathcal{L}(G)$  a regular language? Explain your answer by either giving a reason why it is not or by giving a regular grammar for  $\mathcal{L}(G)$ . (10pt)

**Solution:** .....

$\mathcal{L}(G)$  is regular. We can substitute  $U$  in the first production of  $S$  to obtain

$$S \longrightarrow aaS \mid abS \mid bbS \mid baS \mid \lambda.$$

This can be refined into the following regular grammar.

$$\begin{aligned} S &\longrightarrow aA \mid aB \mid bB \mid bA \mid \lambda \\ A &\longrightarrow aS \\ B &\longrightarrow bS \end{aligned}$$

**More compact solution**

$$\begin{aligned} S &\longrightarrow aT \mid bT \mid \lambda \\ T &\longrightarrow aS \mid bS \end{aligned}$$

$\square$

### 3 Push Down Automata I

**Write your answers on a separate sheet**

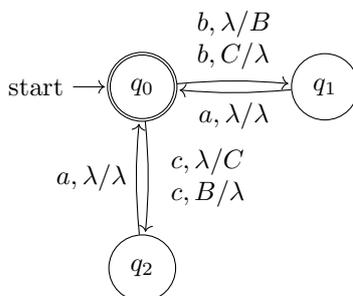
Let  $M$  be the PDA with

$$\begin{aligned} Q &= \{q_0, q_1, q_2\} & \delta(q_0, b, \lambda) &= \{\langle q_1, B \rangle\} \\ \Sigma &= \{a, b, c\} & \delta(q_0, b, C) &= \{\langle q_1, \lambda \rangle\} \\ \Gamma &= \{B, C\} & \delta(q_0, c, \lambda) &= \{\langle q_2, C \rangle\} \\ F &= \{q_0\} & \delta(q_0, c, B) &= \{\langle q_2, \lambda \rangle\} \\ & & \delta(q_1, a, \lambda) &= \{\langle q_0, \lambda \rangle\} \\ & & \delta(q_2, a, \lambda) &= \{\langle q_0, \lambda \rangle\} \end{aligned}$$

a) Draw a state diagram for  $M$ . (5pt)

Solution: .....

$M$  can be drawn as



□

b) Check which of the following words is in  $\mathcal{L}(M)$  and explain your answer:  $abcb$  and  $bac a$ . (5pt)

Solution: .....

- There is no transition from  $q_0$  that reads an  $a$ , so  $abcb \notin \mathcal{L}(M)$ .
- $\langle q_0, bac a, \lambda \rangle \rightarrow \langle q_1, aca, B \rangle \rightarrow \langle q_0, ca, \lambda \rangle \rightarrow \langle q_2, a, C \rangle \rightarrow \langle q_0, \lambda, \lambda \rangle$ .  
We end in an accepting state with an empty stack, so  $bac a \in \mathcal{L}(M)$ .

□

c) Is  $\mathcal{L}((ca)^*(ba)^*) \subseteq \mathcal{L}(M)$ ? Explain your answer. (5pt)

Solution: .....

No,  $ca \in \mathcal{L}((ca)^*(ba)^*)$ , but  $ca \notin \mathcal{L}(M)$ . To see the latter, note that when reading a  $c$ , either a  $B$  must be popped of the stack, or a  $C$  is pushed on the stack. The first is not possible with an empty stack, and after the second the stack is not empty.

Alternatively: if  $w \in \mathcal{L}(M)$  then the number of  $a$ 's and  $b$ 's in  $w$  is equal, following the same argument as above. This does not have to be the case for every  $w \in \mathcal{L}((ca)^*(ba)^*)$ . □

d) Give a precise description of  $\mathcal{L}(M)$  using set notation. (5pt)

Solution: .....

$$\mathcal{L}(M) = \{w \in \mathcal{L}((ba + ca)^*) \mid |w|_b = |w|_c\}.$$

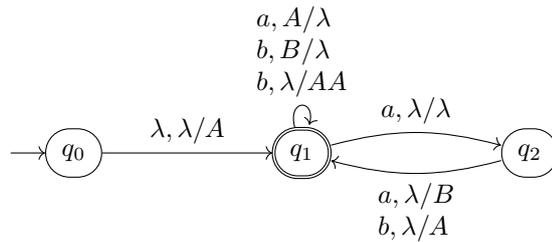
That is, the words with an equal number of  $b$ 's and  $c$ 's, where every  $b$  and every  $c$  is followed by an  $a$ . □

## 4 Push Down Automata II

Write your answers on a separate sheet

- a) i) Let  $A = \{a, b\}$  and let  $L$  be the language  $L = \{w \in A^* \mid |w|_a = 2|w|_b + 1\}$ .  
 Show that  $L$  is context free by giving a PDA that accepts it. (10pt)

Solution: .....



□

- ii) Show that  $aaba$  and  $baaa$  are accepted, by giving the accepting computations. (5pt)

Solution: .....

$$\begin{aligned}
 & q_0, aaba, \lambda \rightarrow q_1, aba, \lambda \rightarrow q_2, ba, \lambda \rightarrow q_1, a, A \rightarrow q_1, \lambda, \lambda \\
 & q_0, baaa, \lambda \rightarrow q_0, aaa, AA \rightarrow q_1, aa, AA \rightarrow q_1, a, A \rightarrow q_1, \lambda, \lambda
 \end{aligned}$$

□

- iii) Show that  $aab$  is not accepted by your PDA. (5pt)

Solution: .....

We have the following possible computations

$$\begin{aligned}
 & (q_0, aab, \lambda) \rightarrow (q_1, aab, a) \\
 & \quad \rightarrow \{(q_1, ab, \lambda), (q_2, ab, a)\} \\
 & \quad \rightarrow \{(q_2, b, \lambda), (q_1, b, ba)\} \\
 & \quad \rightarrow \{(q_1, \lambda, a), (q_1, \lambda, a)\},
 \end{aligned}$$

neither of which ends with an empty stack. Thus  $aab$  is not accepted.

□

- b) Let  $G$  be the grammar on the alphabet  $\{a, b\}$  given as follows.

$$\begin{aligned}
 S & \rightarrow \lambda \mid aX \mid bY \\
 X & \rightarrow bYb \mid bb \\
 Y & \rightarrow aXa \mid aa
 \end{aligned}$$

Construct a PDA that accepts  $\mathcal{L}(G)$ , using the procedure given in the lecture. (10pt)

**Solution:** .....

First we transform the grammar into the form necessary construct a PDA.

$$S \rightarrow \lambda \mid aX \mid bY$$

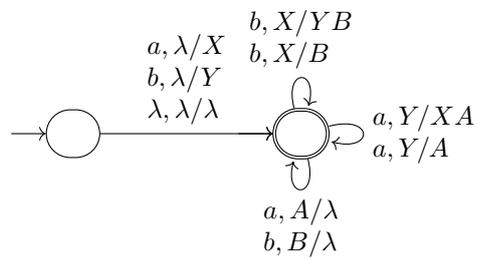
$$X \rightarrow bYB \mid bB$$

$$Y \rightarrow aXA \mid aA$$

$$A \rightarrow a$$

$$B \rightarrow b$$

The resulting PDA is then.



□